Automatic and accurate shadow detection from (potentially) a single image using near-infrared information

Clément Fredembach, Sabine Süsstrunk
EPFL
1015 Lausanne, Switzerland
{clement.fredembach,sabine.susstrunk}@epfl.ch

Abstract

Shadows, due to their prevalence in natural images, are a long studied phenomenon in digital photography and computer vision. Indeed, their presence can be a hindrance for a number of algorithms; accurate detection (and sometimes subsequent removal) of shadows in images is thus of paramount importance.

In this paper, we present a method to detect shadows in a fast and accurate manner. To do so, we employ the inherent sensitivity of digital camera sensors to the near-infrared (NIR) part of the spectrum.

We start by observing that commonly encountered light sources have very distinct spectra in the NIR, and propose that ratios of the colour channels (red, green and blue) to the NIR image gives valuable information about impinging illumination. In addition, we assume that shadows are contained in the darker parts of an image for both visible and NIR. This latter assumption is corroborated by the fact that a number of colorants are transparent to the NIR, thus making parts of the image that are dark in both the visible and NIR prime shadow candidates.

These hypotheses allow for fast, accurate shadow detection in real, complex, scenes, including soft and occlusion shadows. We demonstrate that the process is reliable enough to be performed in-camera on still mosaicked images by simulating a modified colour filter array (CFA) that can simultaneously capture NIR and visible images. Finally, we show that our binary shadow maps can be the input of a matting algorithm to improve their precision in a fully automatic manner.

1 Introduction

A shadow is cast when an object occludes a light source. Due to the difference between the light intensity reaching a shaded region and a directly lit region, shadows often are characterised by conspicuous strong brightness gradients. The stark change between shadow and non-shadow regions is not, however, solely a brightness difference, but a colour one as well; in outdoor scenes, the lit areas are illuminated by the sun and the sky, while the shadow regions are illuminated by the sky alone, thus creating a bluish colour cast.

This property of shadows, coupled to the fact that a certain colour can exist in both lit and shaded objects make shadows problematic in a number of different fields such as tracking [15], object (or people) recognition [27] or white-balancing [12]. Additionally, in image manipulation or compositing, shadows are often unwanted artefacts that are unavoidable due to image capture conditions (e.g., photographs taken in an urban environment), and need to be detected and removed.

While shadow compositing and removal has dramatically improved in recent years [1], [23], [29] automatic shadow detection has not followed suit. Indeed, the best performing methods still require user interaction or image sequences to perform optimally. A review of existing shadow detection methods is provided in Section 2.

The starting point of our approach is the work on joint near-infrared and colour photography presented in [13] and [30] where it has been shown that NIR wavelengths (700-1100 nanometers) could be imaged

Fig. 1. Visible and NIR images of the same scene, captured with identical exposure settings to allow comparison. Notice the darker shadows in the NIR image.

and used in conjunction with a colour image of the same scene for image enhancement. More details about near-infrared/visible photography and its existing applications are provided in Section 3.

Our shadow detection method is, in essence, very simple. We capture two images of the same scene, a standard colour one and a near-infrared one. We make the hypothesis that shadows can be found in the darker parts of the image (both for the visible and NIR), see Fig. 1. Additionally, we assume that the shadows come from a "standard" light difference: shadow areas are illuminated by skylight, incandescent, or fluorescent lights; non-shadow areas are illuminated by sunlight, incandescent, camera flash, or fluorescent lights; shadows and non-shadows areas are illuminated by different lights.

We then proceed by building three different maps: a "dark map" of the visible image, a "dark map" of the NIR image, and a ratio map of NIR to colour. The dark maps will identify the pixels that are shadow candidates, while the ratio map will ensure that these candidates are indeed shadow-related instead of simply being dark objects. Our algorithm is detailed in Section 4.

An important feature of the proposed method is that it can, theoretically, be implemented in-camera on a single image. While currently joint NIR/colour photography requires two images, Lu et al. [20] proposed a colour filter array (CFA) design that can be employed to capture visible and NIR information simultaneously on a single sensor, with little image quality loss. We simulate the output of such a CFA and show in Section 5 that accurate binary shadow masks can be quickly computed on such mosaicked images.

Experiments on a variety of real, complex, scenes demonstrate the validity of our approach in Section 6, while the principle of using our shadow masks in a matting framework to obtain more refined masks (an idea first proposed in [23]) is described in Section 7.

While shadow removal is often a natural continuation of detection, it employs very different assumptions and techniques that are beyond the scope of this paper.

2 The Current State of Shadow Detection

Let us assume a lambertian model of image formation, i.e.,

$$\rho_k = \int_\omega E(\lambda)S(\lambda)Q_k(\lambda)d\lambda \tag{1}$$

Where ρ represents a camera's sensor responses, $E(\lambda)$ is the illuminant spectral power distribution, $S(\lambda)$ is the surface reflectance function, $Q_k(\lambda)$ the sensor sensitivities, λ is the wavelength, and k represent a given channel. Customarily for colour images, $k = \{R; G; B\}$ and $\omega \in [400, 700]$nm, but since our work includes near-infrared information, we will use $k = \{R; G; B; NIR\}$ and $\omega \in [400, 1100]$nm.

Effectively, while (1) predicts the discrete sensor responses arising from a given -spectrally continuous- scene, the reverse problem (finding $E(\lambda)$ and $S(\lambda)$ from the sensor responses) is ill-posed and has an infinite number of solutions. It follows that a given RGB triplet is as representative of, e.g., a bright surface in the shade or a brightly lit dark surface. To accurately detect shadows, one therefore has to obtain additional information about the scene or to make additional assumptions.

Image-based shadow detection algorithms can loosely be classified into two categories depending on the type of additional information they employ: user interaction algorithms and fully automatic methods that have

additional assumptions about the scene. As this paper focuses on still images, video-based methods are not discussed here.

2.1 User interaction methods

A common way to sidestep the difficulty of shadow detection is to assist the detection algorithm with user-supplied information. Leaving aside algorithms that require users to input the entire shadow mask, a number of the most recent and best performing methods incorporate user input to either "seed" or correct the detection process.

In Wu et al. [29], users are asked to submit a quad-map containing shadow, non-shadow, and penumbra regions of similar textures. Simplifications of these user requirements are focused towards reducing the time spent selecting shadow and non-shadow regions. Wu and Tang [28] employ user-supplied context that indicates candidate shadow regions, Arbel and Hel-Or's method [1] allows a shadow mask to be calculated using only a few keypoints, and Shor and Lischinski [23] proposed reducing external information to one user-supplied keypoint per shadow region. Instead of "growing" a shadow region based on a few keypoints, Drew and Reza [7] calculate invariant images based on a few selected patches in the image.

While these detection algorithms are arguably the most accurate ones for still images and can deliver impressive subsequent shadow removal, their requirements are strongly dependent on the complexity of the image. Indeed, an image that contains a single shadow cast on similarly textured background is simple enough to be processed accurately. On the on the other hand, if the scene has a lot of strong material changes in the shadow as well as simultaneous shadow/material changes, e.g., occlusion shadows, it can then become difficult for a user to asses how many keypoints/patches to input. Moreover, even minimal user interaction prevent the detection algorithm to be adapted to a fully automatic workflow, such as in-camera image processing.

2.2 Image-based automatic algorithms

Automatic shadow detection on a single or a couple of images has been addressed in a variety of approaches. Gradient-based methods, where edges are classified as either shadow or material transitions depending on their direction and magnitudes, have been proposed in [17], [25]. Finlayson et al., by assuming Planckian illumination, showed that greyscale illumination-invariant images could be obtained by projecting an image's log chromaticities in an appropriate direction (found by either calibration [10] or entropy minimisation [8]); comparing the edge content of the original image with the edges of the invariant one effectively yields shadow edges. An approach using similar assumptions than [10] and a trichromatic attenuation model has been proposed by Tian et al. [26].

Despite their relatively simple assumptions, these methods often work well and still are state-of-the-art regarding automatic shadow detection from a single image. None of them, however, account for simultaneous material/illumination changes, which can limit their usefulness in more complex scenes.

To minimise the ambiguities induced by simultaneous material and illumination changes, some research has focused towards multi-image (generally two) methods. For instance, flash/no-flash image pairs can be combined to either estimate the illuminant [5], [24] or to remove shadows [6]. The chromagenic illuminant estimation of Finlayson et al. [9] postulated that capturing two images of a given scene, using a broadband coloured filter to capture the second image, and comparing them adequately would produce accurate illumination maps.

In the right context, these multi-image algorithms perform remarkably well since being pixel-based methods they can detect soft shadows as well as occlusion shadows. However, that context is oftentimes limited. Flash cannot illuminate all outdoor shadows, and the chromagenic approach requires image segmentation, an accurate training step, and has a non-negligible complexity due to the fact that every pair of trained illuminants has to be tested. On the other hand, the chromagenic approach is not limited to shadows but can incorporate a number of multi-illuminant scenes.

In previous work [14], we proposed that near-infrared information could be used to identify Planckian illuminants among a training set, and showed that approximate, smooth, illumination maps could easily be computed.

The shadow detection method we detail in this paper builds on this idea, but we show that it can theoretically be performed on a single, mosaicked image, and deliver highly precise shadow maps without requiring a training step.

3 Near-infrared imaging

The visible part of the electromagnetic spectrum, i.e., the part the human visual system (HVS) is sensitive to, comprises wavelengths between 400 and 700 nanometers. The exact definition of near-infrared is often context-dependent, but for digital imaging applications it is defined as the part of the spectrum that lies just beyond the HVS capabilities and up to the sensitivity of the silicon, i.e., from 700 to 1100 nanometers.

The specificity of the near-infrared band has proven to be a source of invaluable information in remote sensing for the detection of water or vegetation [19], spectroscopy for material identification and forgery detection [4], and plant science for the timely detection of crop diseases [21].

In a photographic context, while NIR has been employed since the silver-halide days, it is the advent of high-quality silicon-based sensors (CCD or CMOS) that has allowed camera users to acquire NIR images with relative ease. Such images are commonplace in security applications, such as face recognition [18], video surveillance, or iris biometrics [3].

3.1 Acquiring near-infrared images

Silicon, the photosensitive component of a digital camera's sensor is intrinsically sensitive to the NIR. In fact, its sensitivity is such that a specifically designed filter -often called a hot mirror- has to be placed in front of the sensor to prevent NIR contamination of the visible images. This filter is necessary for both monochromatic and colour cameras because the filter colorants employed to create the colour filter array (CFA) are also transparent to the NIR. Figure 2 shows the sensitivity of silicon as well as the transmittance of usual CFA filters. It follows that removing the hot mirror from a camera gives a user the potential of acquiring both visible and NIR images.

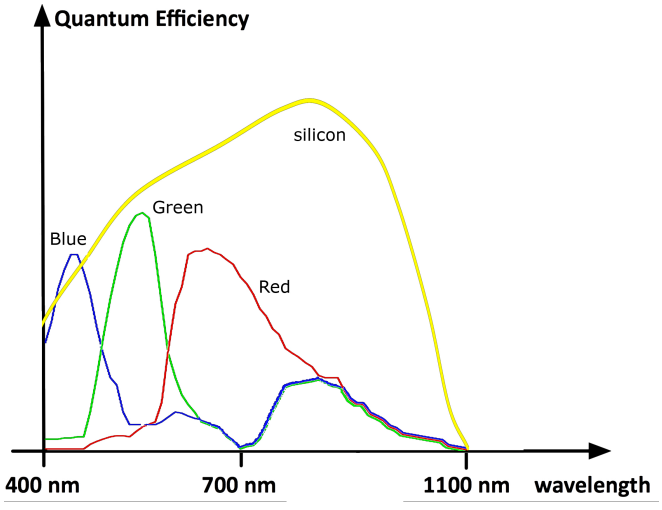

Fig. 2. The sensitivity of silicon and the transmittance of the red, green, and blue filters of a classical Bayer CFA represented as quantum efficiency.

To obtain relevant information, one needs to separate the visible and NIR signals that reach the imaging sensor. This can be done by a computational photography set-up: a beamsplitter is used to separate visible wavelengths from NIR ones and two (or more) cameras are used to capture a portion of the spectrum only [30]. Another, single camera, approach consists in capturing two images of the same scene: A colour image

(by affixing a hot-mirror in front of the lens) and a NIR image (placing a visible light blocking filter[1] in front of the lens). Here, we use the latter approach.

As the notion of colour is ill-defined when dealing with data outside of the HVS sensitivities, NIR images often are represented as greyscale intensity; a pair of colour/NIR image of the same scene is shown in Fig. 1. Importantly for this work, the NIR and visible images are captured with *identical* exposure settings to allow for a meaningful image comparison.

3.2 Applications of joint colour/NIR imaging

Recent research has shown that colour and NIR images of the same scene could be used for image enhancement. Zhang et al. in [30] demonstrated a 2-camera setup to capture NIR and colour images simultaneously. Their work, aimed at high dynamic range imaging, combined NIR and visible information to produce an improved colour rendering for HDR scenes. Additionally, they showed that the gradient distribution of NIR images was similar to that of visible images, an important step for realistic image fusion.

Similarly, in [13] Fredembach and Süsstrunk proposed that NIR be treated as a luminance image. Combining NIR with the chrominance of colour images would yield colour images that had a greater local contrast and less haze. This effectively showed that NIR and visible images have little correlation in the low-frequencies, but a significant one in the high-frequencies (i.e., edges are mostly located in the same part of both NIR and visible luminance). Figure 3 illustrates the enhancement that can be obtained by these two approaches.

Fig. 3. Original colour images and their NIR-enhanced version according to the method of [30] (top) and [13] (bottom). The enhancements are better appreciated as large images on screen.

Application specific algorithms that use joint near-infrared and visible information have been proposed for color image dehazing [22], skin smoothing [11], and Planckian illuminant estimation [14].

4 Hypotheses for shadow detection

In the event where two (significantly) spectrally different lights are present in a scene, the shadows that will be created will differ from the non-shadow regions in terms of brightness and colour. This variation will depend on the scene's illumination but can be realistically reduced to one of the following cases: skylight/sunlight, fluorescent/incandescent, flash/incandescent, and flash/fluorescent. The outdoor shadows case is by far the most common and problematic for computer vision.

1. e.g., a Wratten 97 or equivalent

We formulate here two hypotheses. First, shadows are generally found in the dark parts of an image, be it colour or NIR; these dark pixels are shadow *candidates*. Second, all the considered illuminants in the shadow formation process have a very distinct behaviour in the NIR (see Fig. 5) that can be employed to assess the shadow candidate pixels.

4.1 Shadows are dark

Saying that shadows are dark may sound like a truism, but is not necessarily so. In the image formation process, the defining property of shadows is that they are *darker than their surroundings*. Equating shadows with dark is an often used but misguided shortcut because all dark objects will become potential shadows. Nonetheless, NIR possesses some important properties that make the shadow candidate selection more precise.

Spectral studies of natural and man-made surfaces and colorants [2], [4], [19] show that in general:

$$\int_{VIS} S(\lambda)Q_{R,G,B}(\lambda)d\lambda < \int_{NIR} S(\lambda)Q_{NIR}(\lambda)d\lambda \tag{2}$$

i.e., the reflectance times quantum efficiency in the NIR is greater than in any of the RGB channels. Two notable exceptions are water, which has an absorption band in the NIR, and carbon black, a colorant often encountered in black plastic objects where NIR and the visible have a similar, almost zero reflectance. The implication of (2) is quite profound as it permits to disambiguate a number of otherwise problematic dark objects/surfaces.

The dark maps we will use to identify shadow pixel candidates are constructed as follows. Let ρ_k be the normalised sensor responses from (1), i.e.,

$$0 \leq \rho_k \leq 1 \; k \in \{R; G; B; NIR\} \tag{3}$$

We first create a *brightness* image L from the colour image simply by calculating the pixels' norm in an RGB cube:

$$L = \frac{\rho_R + \rho_G + \rho_B}{3} \tag{4}$$

Temporary dark maps D_{VIS} and D_{NIR} are computed by:

$$D_{VIS} = 1 - L \; ; \; D_{NIR} = 1 - \rho_{NIR} \tag{5}$$

Before calculating the entire dark map, we make the observation that the relationship between D_{VIS}, D_{NIR}, and the presence of shadows is certainly not linear. Indeed, shadows are never going to be found in the highlights of an image. To obtain a more accurate model, we must apply the physical knowledge that has been employed since the days of silver-halide photography.

Cameras are designed to accurately capture midtones. Shadows and highlights are usually compressed at each end of an S-shaped curve. Incidentally, this behaviour is one we are looking for; by compressing shadows we are effectively marking fewer but better controlled pixels as shadow candidates. Our global dark map D is thus,

$$D = f(D_{VIS})f(D_{NIR}) \tag{6}$$

where

$$f(x) = \frac{1}{1 + e^{-a(x-b)}} \tag{7}$$

a and b are parameters that we have chosen to be 10.0 and 0.5, respectively.

The multiplication of the two maps in (6) comes from the fact that both $f(D_{VIS})$ and $f(D_{NIR})$ lie between 0 and 1. While they are not, formally speaking, probability functions, they can be considered as such since darkness in *both* visible and NIR images is a condition of shadow presence.

Figure 4 shows the importance NIR images can have in disambiguation. A black jacket lies partly in the shade, but since its colorant is transparent to the NIR wavelengths, it appears bright in the NIR image. Consequently, the total dark map will be much more shadow oriented than the visible one.

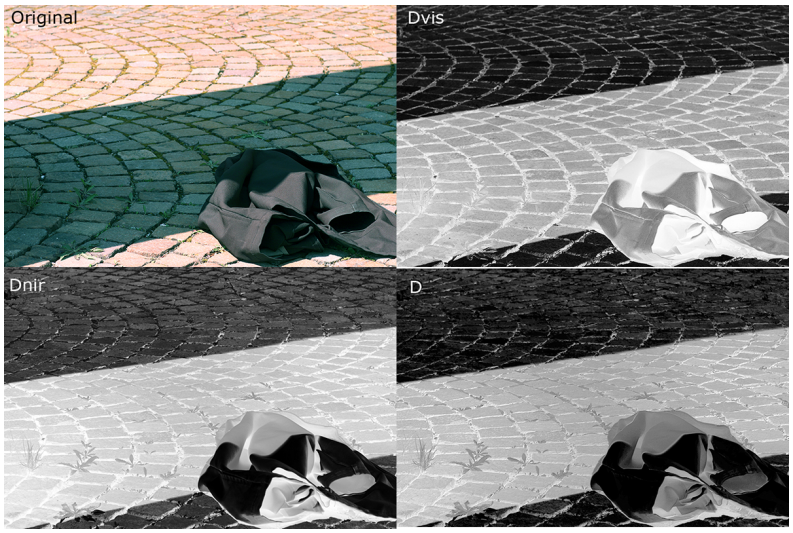

Fig. 4. A colour image, D_{VIS}, D_{NIR}, and D (bright pixels = likely shadow). While the presence of a black object in the scene confounds D_{VIS} , D is quite accurate.

4.2 Colour to NIR ratios

The key reason of using colour to NIR ratios is best illustrated by Fig. 5. If one focuses towards the 700-1100nm (NIR) region, the difference between the visible and NIR bands is striking. For the sake of simplicity, let us consider the case of outdoor shadows, i.e., the difference between skylight and sunlight. However, the following derivation is valid for *all* the considered light combinations that create shadows (e.g., incandescent/fluorescent, flash/fluorescent, daylight/indoor light, flash/incandescent).

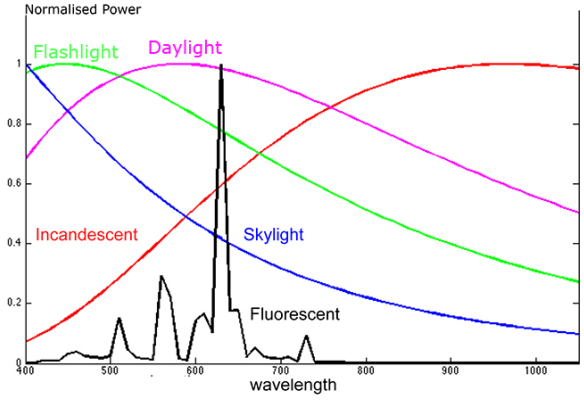

Fig. 5. Measured spectra of typical shadow-creating illuminants. Note the large differences in the NIR band compared to the visible one.

Irrespectively of the difference in light intensity due to occluding objects, we see that skylight emits very little in the NIR, while sunlight actually emits approximatively as much energy in the NIR than in the visible band. It follows that we can write

$$\int_{VIS} E_{sky}(\lambda)Q_k(\lambda) \quad >> \quad \int_{NIR} E_{sky}(\lambda)Q_k(\lambda) \tag{8}$$

$$\int_{VIS} E_{sun}(\lambda)Q_k(\lambda) \quad \approx \quad \int_{NIR} E_{sun}(\lambda)Q_k(\lambda) \tag{9}$$

Using (8) and (9) in conjunction with (2) implies that we can calculate image ratios that will have a

significant impact on shadow detection. Specifically, we compute a ratio image F as:

$$F = \frac{1}{2}\min(\max_k(F_k), 2) \quad (10)$$

$$F_k = \frac{\rho_k}{\rho_{NIR}};\ k = \{R, G, B\} \quad (11)$$

The rationale of (10)-(11) is that ratios can take very large values but the main point of interest is whether they are closer to 1 or to 0. The max operator is there because scene reflectances can often have very low values in one or two of the colour channels, although rarely in all three (see [30]). From (2) we can deduce that for sunlit regions $F < 1$, while for skylit regions $F > 1$ because the difference in illumination far outstrips the difference in reflectances in the shade. Constraining F to be in the [0-1] range also allows a more meaningful comparison with D.

The ratio image, illustrated in Fig. 6, is not, however, sufficient to detect shadows due to the potential variability of reflectances in both the visible and NIR part of the spectrum.

Fig. 6. Colour image, near-infrared image, and the ratio image F that already outlines the shadow

To obtain the final shadow map M, we have to take into account both our dark hypothesis and the ratio image. We know that D contains all the possible shadow pixels but can also include dark objects. On the other hand, if both the visible and NIR pixel values of a given objects are dark, then F will adequately be able to discriminate being an actual shadow and a dark object. Since both D and F have comparable values, we simply write:

$$M = (1 - D)(1 - F) \quad (12)$$

The values of M will be in the range [0-1].

To obtain a fixed shadow mask, we have to binarise M, which we do according to the work of Shor and Lischinski [23]. Basically, we compute the histogram of M and calculate the location of its first valley. Let θ be this value. The shadow value of each pixel x in the image is then given by:

$$M_{bin}(x) = \begin{cases} 1 & \text{if } M(x) \leq \theta \\ 0 & \text{otherwise} \end{cases}$$

The outcome of this thresholding operation is illustrated in Fig. 7, where an accurate binary mask is obtained.

5 In-camera shadow detection

Up to now, we have illustrated the working of our shadow detection algorithm by assuming that full resolution colour and NIR images are available for each scene. Capturing such images does, however, require either a multi-camera or multi-exposure setup.

Though little investigated so far, combining NIR and visible images for digital photography and computer vision has shown great potential in a number of applications. There is thus research towards enabling simultaneous capture of both visible and NIR radiation with a single sensor, such as the modified CFA presented by Lu et al. in [20]. Considering the low complexity of our method, it can realistically be implemented in-camera to provide assistance where shadow detection is beneficial, such as in white-balancing or automatic image enhancement.

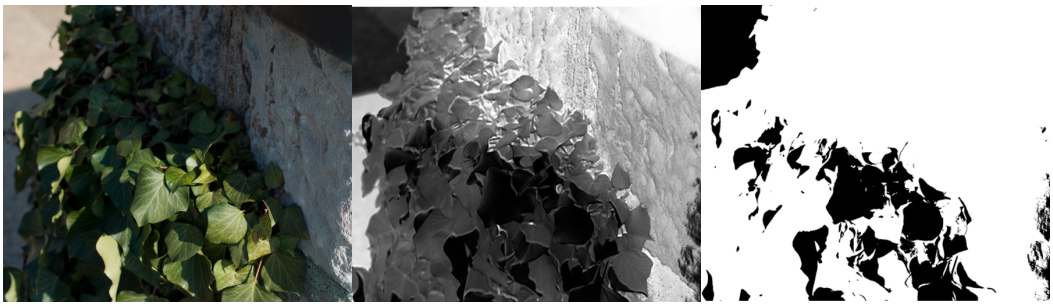

Fig. 7. A colour image, the shadow map M (brighter implies a greater likelihood of shadows) and its binary version M_{bin}

To assess the robustness of our method, we simulate joint NIR/visible capture as follows: We take, successively, an NIR and visible image of the same scene in camera RAW format. We then replace one of the green values of the visible image (the Bayer CFA has 2 green pixels for every red and blue) by its corresponding NIR value. Finally, from this modified mosaicked image, we obtain four equally sized images (R, G, B, NIR) by selecting only the relevant CFA pixels, i.e., no interpolation is made, only downsampling. This procedure is illustrated in Fig. 8

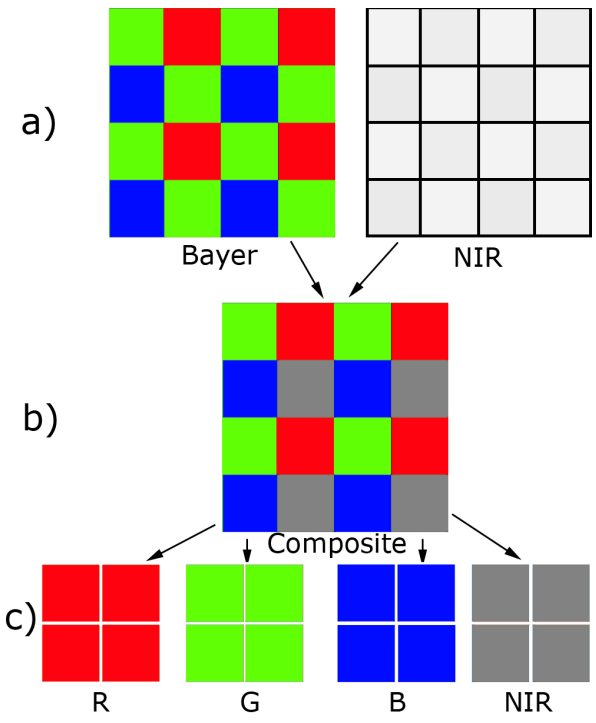

Fig. 8. Illustration of our simulation workflow: From the two original images **(a)**, we replace a green pixel by a NIR one at the same location **(b)**. We obtain the R, G, B, and NIR images on which shadow detection is subsequently performed by selecting only the relevant pixels from the composite mosaic **(c)**.

Downsampling in such a fashion, instead of demosaicking, will make the shadow map a bit less precise around edges, but this is a low concern. Indeed, most subsequent applications of shadow detection (apart from shadow removal), do not require a pixel-level degree of precision. Regarding shadow removal, most recent work [1], [23] shows that a binary mask was not sufficient for graphics-quality shadow removal, so a

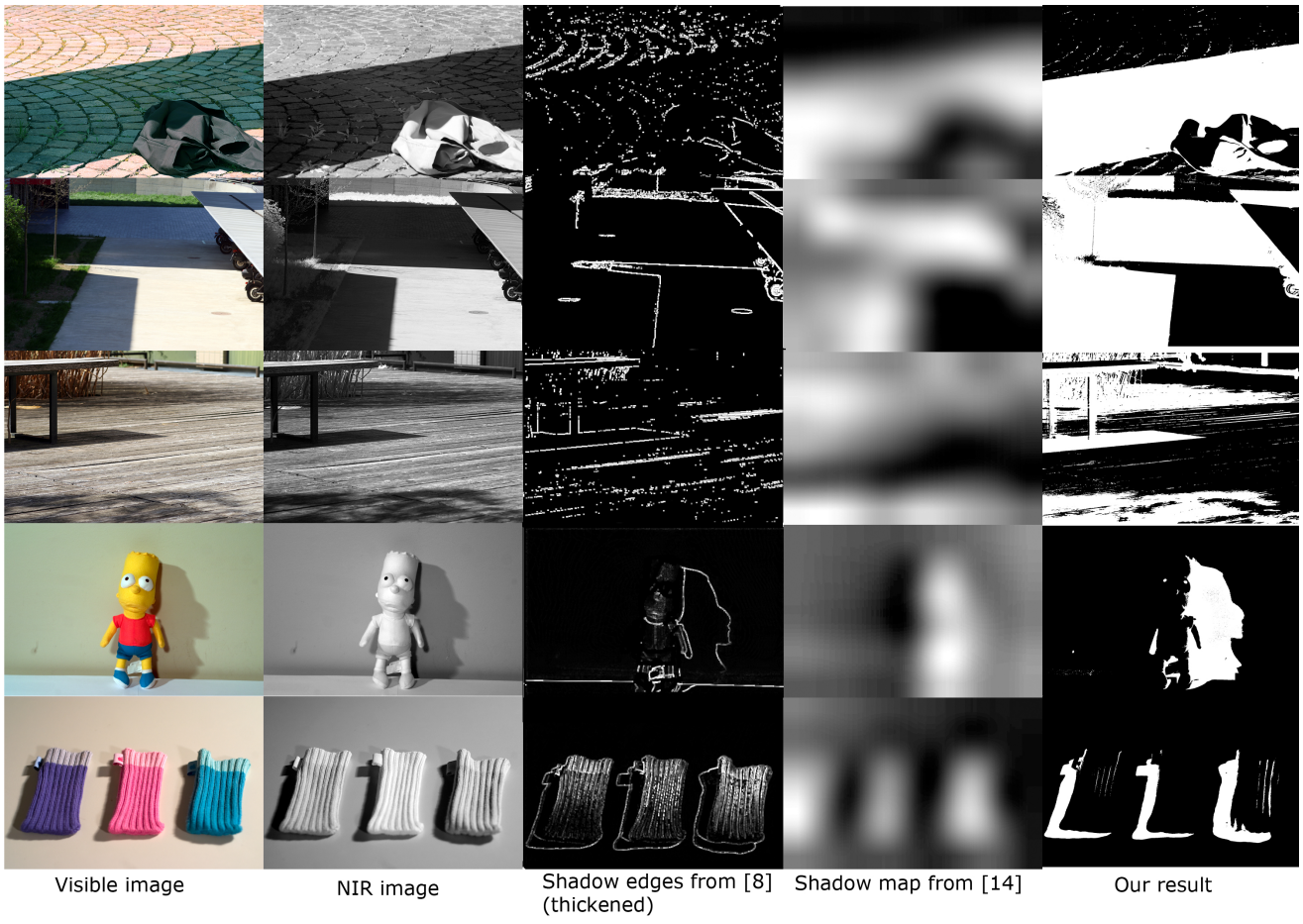

Fig. 9. Input images and results from Finlayson et al. intrinsic images [8] (edges thickened for better visibility in this figure), the interpolated map from [14], and our results. The results from our algorithm are very accurate.

difference of a few pixels in a binary map is not critical. Previous removal methods such as [8], [10] dilate the map prior to reintegration.

Computation-wise, our entire shadow detection procedure consists in a pixel-wise division and multiplication, and a histogram-based thresholding. The binary shadow map can thus be computed at frame-rate for in-camera processing.

6 Results

Figure 9 presents the shadow detection results of our proposed method as well as a comparison with the prior art of [8] and [14]. The first three images are captured outdoors, while the last two illustrate the flash/fluorescent case. We see that the shadow edge detection of [8] performs well on dark shadows cast against a relatively smooth background (second and fourth row images). The method gets less reliable when more textured elements are introduced (first and fifth row images), when softer shadows are present (third row image), or in the presence of coincidental material and illuminant edges(parts of the roof in the second image, open edges in the fourth). The soft, interpolated, maps of [14] are more resilient to texture because of the downsampling and interpolation, but are imprecise, especially when small shadows are present. Generally, they do not offer a sufficient amount of detail for applications such as shadow removal.

On the other hand, the proposed method, performed on *still mosaicked* images, produces accurate shadow maps even in difficult conditions (e.g., colour cast and dark jacket in the first image, soft shadows, textured

background). Slight detection mistakes occur when the images are underexposed (the bikes in the second row image and the doll's eyes in the fourth) or when the shadows are very faint (top of the socks in the last row image). These situations, however, also negatively impact the performance of all other automatic methods.

7 Binary masks as preprocessing for matting

A possibility to further improve our shadow masks is to use a matting algorithm. This idea has been successfully used before in [23] and we apply here their framework of eroding our binary masks prior to treating them as light/dark scribbles in Levin et al.'s matting algorithm [16].

While this post-processing significantly increases the complexity of shadow detection, it also yields much improved masks, shown in Fig. 10, that can then be employed for shadow removal. Moreover, no user input is required.

Fig. 10. The visible image, reduced binary shadow mask (the input to the matting algorithm) and shadow mask obtained by the method of Levin et al. [16]

8 Conclusions

While digital imaging sensors are inherently sensitive to both visible and near-infrared wavelengths, combining this information for meaningful image processing is still little studied.

In this work, we have presented a simple algorithm for accurate, fully automatic, shadow detection in images using physical properties of the near-infrared band, such as generally higher reflectances and very marked differences in illuminants' spectral power distribution. By making two simple hypotheses, we have devised a "dark map" and a "ratio map" that can be combined and thresholded to yield a binary shadow map.

Because the proposed algorithm is simple and robust, we proposed it could be used in-camera. To that effect, we have simulated the output of a CFA that allows for joint NIR/visible image capture and have shown that our method could still produce accurate maps that detect occlusion shadows as well as soft shadows.

Finally, we demonstrated that our binary shadow masks could be the input of a matting algorithm in order to produce maps that can be employed in subsequent shadow removal.

References

[1] E. Arbel and H. Hel-Or. Texture-preserving shadow removal in color images containing curved surfaces. In *Proc. of IEEE Computer Vision and Pattern Recognition*, pages 1–8, 2007.

[2] M. Blue and S. Perkowitz. Reflectivity of common materials in the submillimeter region. *IEEE Trans. on microwave theory and techniques*, 6:491–493, 1977.

[3] K. Boyer, K. Hollingsworth, and P. Flynn. Image understanding for iris biometrics: A survey. *Computer vision and image understanding*, 110:281–307, 2008.

[4] D. Burns and E. Ciurczak. *Handbook of near-infrared analysis*. CRC Press, 2007.

[5] J. DiCarlo, F. Xiao, and B. Wandell. Illuminating illumination. In *Proc. of IS&T/SID 9th Color Imaging Conference*, pages 27–34, 2001.

[6] M. Drew, C. Lu, and G. Finlayson. Removing shadows using flash/no-flash image edges. In *Proc. IEEE International conference on Multimedia Expo*, pages 1–8, 2006.

[7] M. S. Drew and H. Reza. Sharpening from shadows: sensor transforms for removing shadows from a single image. In *Proc. of IS&T/SID 17th Color Imaging Conference*, pages 267–271, 2009.

[8] G. Finlayson, M. Drew, and C. Lu. Entropy minimisation for shadow removal. *International Journal of Computer Vision*, 85:35–57, 2009.

[9] G. Finlayson, C. Fredembach, and M. Drew. Detecting illumination in images. In *Proc. IEEE International Conference on Computer Vision*, pages 1–8, 2007.

[10] G. Finlayson, S. Hordley, C. Lu, and M. Drew. On the removal of shadows from images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 28:59–68, 2006.

[11] C. Fredembach, N. Barbuscia, and S. Süsstrunk. Combining near-infrared and visible images for realistic skin smoothing. In *Proc. of IS&T/SID 17th Color Imaging Conference*, pages 100–106, 2009.

[12] C. Fredembach and G. Finlayson. The bright-chromagenic algorithm for illuminant estimation. *Journal of the IS& T*, 52:1–11, 2008.

[13] C. Fredembach and S. Süsstrunk. Colouring the near-infrared. In *Proc. of IS&T/SID 16th Color Imaging Conference*, pages 176–182, 2008.

[14] C. Fredembach and S. Süsstrunk. Illuminant estimation and detection using near infrared. In *IS&T/SPIE Electronic Imaging, Digital Photography V*, volume 7250, 2009.

[15] H. Jiang and M. Drew. Tracking objects with shadows. In *CME03: International Conference on Multimedia and Expo,*, pages 100–105, 2003.

[16] A. Levin, D. Lischinski, and Y. Weiss. A closed-form solution to natural image matting. *IEEE Trans. on Pattern analysis and Machine Intelligence*, 30:1–15, 2008.

[17] M. Levine and J. Bhattacharyya. Removing shadows. *Pattern Recognition Letters*, 26:251–265, 2005.

[18] S. Li, R. Chu, S. Liao, and L. Zhang. Illumination invariant face recognition using near-infrared images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 29:627–639, 2007.

[19] T. Lilesand and R. Kiefer. *Remote Sensing and Image Interpretation*. Wiley and Sons, 1994.

[20] Y. Lu, C. Fredembach, M. Vetterli, and S. Süsstrunk. Designing color filter arrays for the joint capture of visible and near-infrared images. In *Proc. of IEEE International Conference on Image Processing*, pages 1–4, 2009.

[21] N. Milton, B. Eisworth, and C. Ager. Effect of phosphorus deficiency on the spectral reflectance and morphology of soybean plants. *Remote sensing of the environment*, 3:121–127, 1991.

[22] L. Schaul, C. Fredembach, and S. Süsstrunk. Color image dehazing using near-infrared. In *Proc. IEEE International Conference on Image Processing*, pages 1–4, 2009.

[23] Y. Shor and D. Lischinski. The shadow meets the mask: Pyramid-based shadow removal. *Computer Graphics Forum*, 27:577–586, 2008.

[24] R. Szeliski, S. Avidan, and P. Anandan. Layer extraction from multiple images containing reflections and transparency. In *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 246–253, 2000.

[25] M. Tappen, W. Freeman, and E. Adelson. Recovering intrinsic images from a single image. In *Proc. of the Advances in Neural Information Processing Systems (NIPS)*, pages 1343–1350, 2003.

[26] J. Tian, J. Sun, and Y. Tang. Tricolor attenuation model for shadow detection. *IEEE Trans. on Image Processing*, 18:2366–2363, 2009.

[27] M. A. Turk and A. P. Pentland. Face recognition using eigenfaces. In *Proc. of IEEE Computer Vision and Pattern Recognition*, pages 586–591, 1991.

[28] T.-P. Wu and C.-K. Tang. A bayesian approach for shadow extraction from a single image. In *Proc. of IEEE 10th International Conference on Computer Vision*, pages 480–487, 2005.

[29] T.-P. Wu, C.-K. Tang, M. S. Brown, and H.-Y. Shum. Natural shadow matting. *ACM Transaction on Graphics*, 26, 2007.

[30] X. Zhang, T. Sim, and X. Miao. Enhancing photographs with near infra-red images. In *Proc. of IEEE Computer Vision and Pattern Recognition*, pages 1–8, 2008.